Optical Character Recognition

Minju Kim

2022.03.19.Sat

What is Optical Character Recognition(OCR)?

- Optical Character Recognition (OCR) is the process of detecting and reading text in images through computer vision.

- OCR technology uses pattern recognition and feature detection methodology to convert virtually any kind of images containing written text (typed, handwritten, or printed) into machine-readable text data.
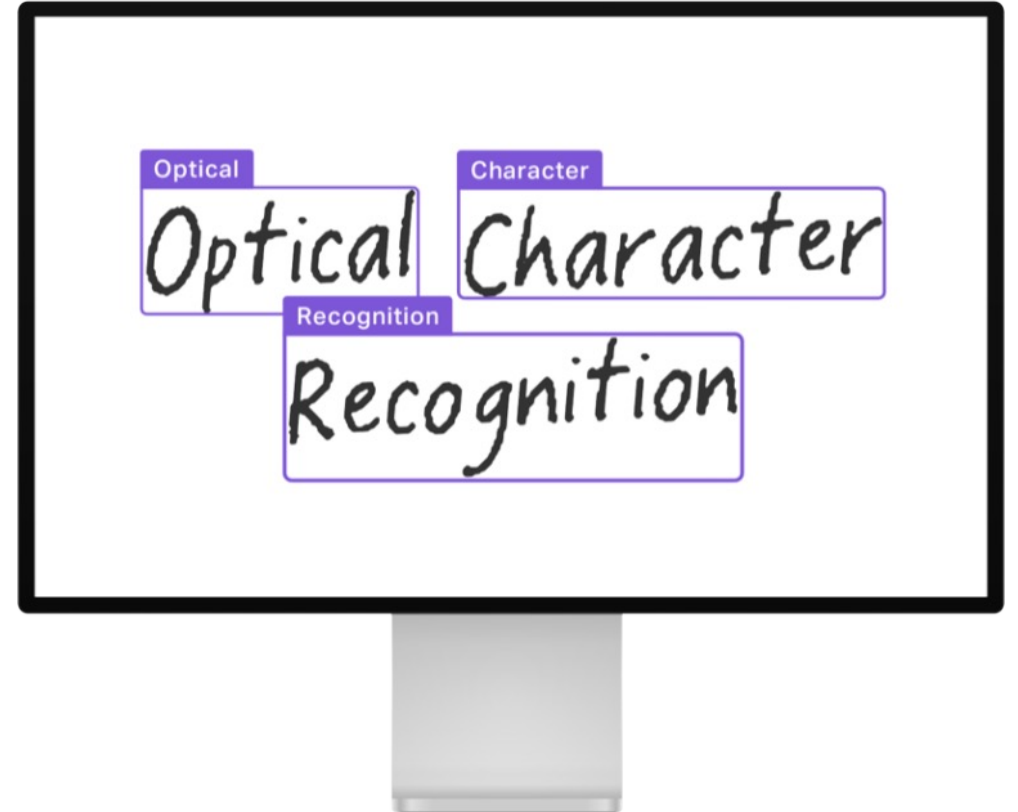
How OCR works?

1. Region Proposal

+
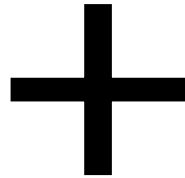
2. Language Processing

- The first stage for OCR involves the detection of textual regions from the image. This is achieved by using convolutional models that detect segments of text and enclose them in bounding boxes.

- NLP-based networks like RNNs and Transformers work to extract information captured in these regions and construct meaningful sentences based on features fed from the CNN layers.

What makes OCR work well?

1. **Denoise Input Data**
   - Data fed to the model should be properly denoised to prevent non-textual regions from being proposed as text. Denoising can be done in several ways, with Gaussian blurring being the most popular. Additive white noise can also be removed with the help of an auxiliary autoencoder network.
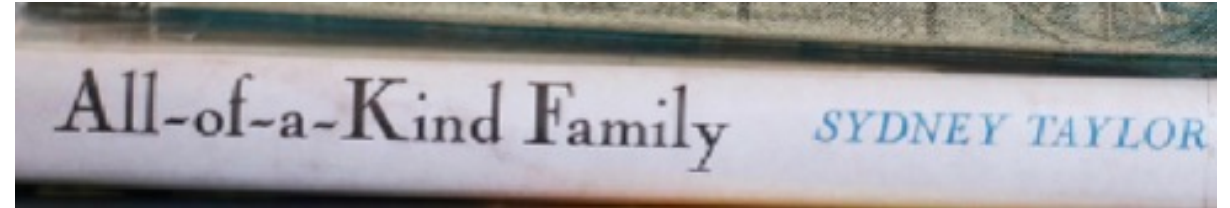
2. **Improve Image Contrast**
   - Image contrast plays a great role in helping the neural network discern textual regions from non-textual ones. Increasing contrast differences between text and background helps the OCR model perform much better.

All-of-a-Kind Family

All-of-a-Kind Family

- Using Median blurs, which compute the median of neighboring pixels to replace the current filter.
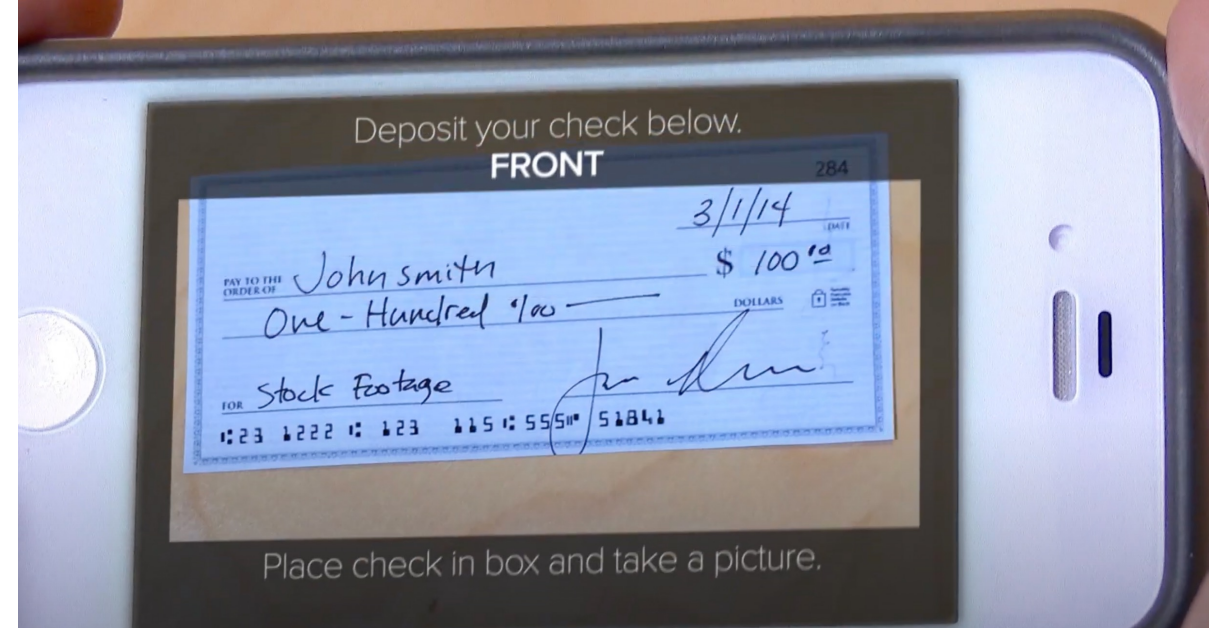
OCR Applications
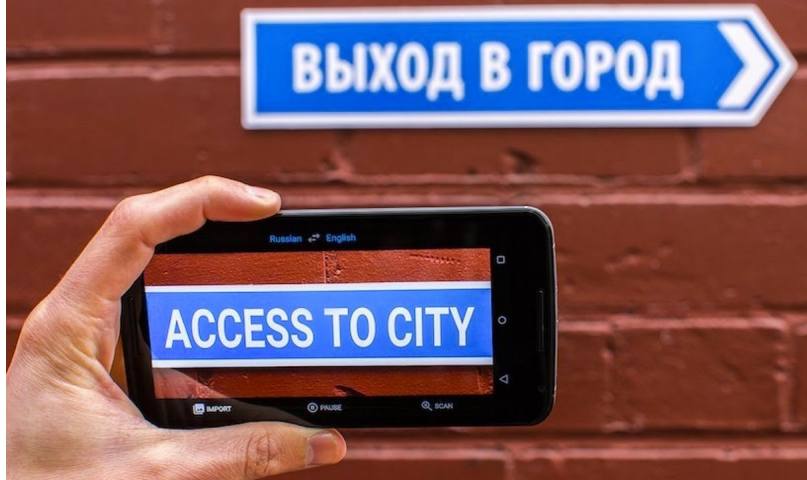
Text Translation
- OCR helps translate written text to a plethora of languages, which increases the accessibility to documents and helps bridge the language gap.
- Text translation forms an important part of OCR, particularly for scene text recognition and evaluation. Translation modules stacked onto the output from an OCR system can help international tourists understand documents and billboards in different languages.

OCR Applications

Archives and digital libraries creation

- OCR helps in digitizing old documents, thereby making preservation extremely easy and secure.

Banking Industries

- OCR is used for processing handwritten checks.

References

- https://viso.ai/computer-vision/optical-character-recognition-ocr/
- https://www.v7labs.com/blog/ocr-guide
- https://www.linkedin.com/pulse/gentle-introduction-ocr-google-translation-susheel-aakulu/
- https://cloud.google.com/vision/docs/ocr
- https://medium.com/analytics-vidhya/a-hitchhikers-guide-to-ocr-8b869f4e3743
- https://www.youtube.com/watch?v=65_LDfZ_hdY

Thank you